(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
28 March 2002 (28.03.2002)

PCT

(10) International Publication Number
WO 02/25233 A2

(51) International Patent Classification[7]: G01J 3/00

(21) International Application Number: PCT/US01/25706

(22) International Filing Date: 16 August 2001 (16.08.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
09/664,973 18 September 2000 (18.09.2000) US

(71) Applicant: INSTRUMENTATION METRICS, INC. [US/US]; 7470 West Chandler Blvd., Chandler, AZ 85226 (US).

(72) Inventors: HAZEN, Kevin, H.; 1534 W. Islandja Drive, Gilbert, AZ 85233 (US). BLANK, Thomas, B.; 2922 E. Tulsa Street, Chandler, AZ 85225 (US). MONFRE, Stephen, L.; 1289 E. Palo Blanco Way, Gilbert, AZ 85296 (US). RUCHTI, Timothy, L.; 1501 West Sea Haze Drive, Gilbert, AZ 85233 (US).

(74) Agents: GLENN, Michael et al.; Glenn Patent Group, Ste. L., 3475 Edison Way, Menlo Park, CA 94025 (US).

(81) Designated States *(national)*: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 02/25233 A2

(54) Title: A METHOD OF CHARACTERIZING SPECTROMETER INSTRUMENTS AND PROVIDING CALIBRATION MODELS TO COMPENSATE FOR INSTRUMENT VARIATION

(57) Abstract: Spectrometer instruments are characterized by classifying their spectra into previously defined clusters. The spectra are mapped to the clusters and a classification is made based on similarity of extracted spectral features to one of the previously defined clusters. Calibration models for each cluster are provided to compensate for instrumental variation. Calibration models are provided either by transferring a master calibration to slave calibrations or by calculating a separate calibration for each cluster. A simplified method of calibration transfer maps clusters to eachother, so that a calibration transferred between clusters models only the difference between the two clusters, substantially reducing the complexity of the model.

BNSDOCID: <WO 0225233A2>

A METHOD OF CHARACTERIZING SPECTROMETER INSTRUMENTS AND PROVIDING CALIBRATION MODELS TO COMPENSATE FOR INSTRUMENT VARIATION

BACKGROUND OF THE INVENTION

FIELD OF THE INVENTION

The invention relates to variation in spectrometer instruments. More particularly the invention relates to characterizing spectrometer instruments by classifying their spectral responses into a limited number of clusters and developing calibration transfer models between clusters that compensate for instrument variations.

DESCRIPTION OF THE PRIOR ART

Many of the analytical applications for spectrometers require calibration data sets that are time-consuming and expensive to create. Typically, these calibrations are highly specific. For example, apparently identical instruments produced by the same manufacturer may exhibit minor instrument variations; such variations may be seen when one instrument is built with a component that varies slightly from the same component in another instrument. In addition, a calibration set for an instrument produced by one manufacturer is generally not suitable for a similar instrument produced by another manufacturer. Furthermore, repairs to a single instrument can cause the instrument's spectral response to vary. As an instrument ages, it's spectral response may change. An instrument's spectral response may vary according to fluctuations in the operating environment. In applications requiring analysis of very low concentration analytes, non-invasive blood glucose prediction, for example, even minor instrument variation can introduce an unacceptable degree of error into the analysis. Providing another calibration model that takes the instrument's current spectral response into account can compensate for instrument variation. However, development of new calibration models is time-consuming, labor-intensive and costly.

In the development of spectroscopy-based analyzers for biomedical applications, there is a need for production of thousands to as many as millions of analyzers for a specific application. No methodology exists for providing calibrations for large numbers of instruments quickly and inexpensively.

Therefore, efforts have been directed at transferring calibrations from one analyzer to another. See, for example, E. Bouveresse, C. Hartmann, D. Massart, I. Last, K. Prebble, *Standardization of near-infrared spectrometric instruments*, Anal. Chem., vol. 68, pp. 982-990 (1996) and M. Defernez, R. Wilson, *Infrared spectroscopy: instrumental factors affecting the long-term validity of chemometric models*, Anal. Chem., vol. 69, pp. 1288-1294 (1997), and E. Bouveresse, D. Massart, P. Dardenne, *Calibration transfer across near-infrared spectrometric instruments using Shenk's algorithm: effects of different standardization samples*, Analytica Chimica Acta, vol. 297, pp. 405-416, (1994) and Y. Wang, D. Veltkamp, B. Kowalski, *Multivariate instrument* calibration, Anal. Chem., vol. 63, pp. 2750-2756 (1991).

Most of the reported methods of calibration transfer have been applied in situations involving high-concentration analytes, wherein the signal-to-noise ratio is high. Because these currently known methods act as a smoothing function when transferring calibrations, they degrade the signal to noise that can be observed, thus hindering analysis of low concentration analytes. Additional problems of changes in resolution or bandwith across time or between instruments have not been addressed.

Furthermore, the currently known methods have only been successfully applied in situations involving a small number of instruments. The reported methods are not capable of modeling the complexity encountered when large numbers of instruments are involved.

A need exists for the calibration of large numbers of analyzers. It would be desirable to provide a means of reducing the complexity inherent in the transfer of calibrations to large numbers of analyzers. It would also be advantageous to provide a means of transferring calibrations without significant degradation of the signal-to-noise ratio, rendering calibration transfer practical in analysis of low concentration analytes.

SUMMARY OF THE INVENTION

The invention provides a method of characterizing large numbers of spectrometer instruments by classifying their spectra into a limited number of previously defined clusters according to similarity of spectral features and

performance characteristics. The method may also be used to track variation over time within a single instrument. The spectral features used for classification may be related to known instrument parameters, or they may be abstract features derived using a variety of computational methods. The clusters are defined in advance based on an exemplar data set, using either supervised or unsupervised methods. Calibration models for each cluster compensate for instrument variation, either across instruments or across time within a single instrument.

In a preferred embodiment of the invention, calibration models are provided using a method of calibration transfer wherein the clusters are mapped to each other, so that a calibration transferred from one cluster to another need only model the difference between the two clusters. In an alternate embodiment, a different calibration is separately calculated for each cluster. In either embodiment, the number of calibration transfers is significantly reduced since each cluster represents a multitude of instruments.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 provides a block schematic diagram of a method of characterizing spectrometer instruments by clustering according to spectral features, according to the invention;

Figure 2 provides a flow diagram of a method of generating calibration models by transferring a master calibration to slave calibration models, according to the invention;

Figures 3A and 3B illustrate decreases in spectral cutoff related to decreases in detector temperature in a spectrometer instrument; and

Figure 4 illustrates changes in light throughput related to changes in source temperature in a spectrometer instrument.

DETAILED DESCRIPTION

The invention provides a method of characterizing spectrometer instruments by classifying spectral responses from a large number of spectrometers (analyzers) into a limited number of previously defined clusters. The clusters constitute

groups that are defined based on the similarity of specific features. Grouping of instruments according to spectral features and performance characteristics reduces the spectral variation between instruments in a given group. Therefore, spectral measurements corresponding to an individual cluster are more homogeneous than those from the entire group of analyzers. Calibrations created for each cluster may then be utilized for subsequent analysis. The calibration created for a cluster will have instrument variations included in the model that mirror variations resulting from the particular analyzer employed for the subsequent analysis. Calibration models specific to the clusters are expected to be less complex and have an improved level of accuracy. Therefore, multivariate analysis will require fewer factors to model instrument variation resulting in earlier factors modeling the sample and earlier factors modeling small absorbing analytes. The net result is that fewer factors are required for sample analysis and a more robust algorithm is created.

In one embodiment of the invention, a separate calibration data set is required for every cluster, requiring a large number of samples in a calibration set for every cluster. In the preferred embodiment, sufficient samples for a calibration model are only required in a single cluster. Principal features defining that cluster and other clusters are determined, and the clusters are mapped to each other, revealing the specific differences from one cluster to another. Subsequently, the calibration from the initial cluster is transferred to another cluster based upon the specific differences between clusters.

The parent application to the current application, S. Malin and T. Ruchti, *An intelligent system for noninvasive blood analyte prediction*, U.S. Patent Application Ser. No. 09/359,191 (July 22, 1999), provides a detailed description of a method of classifying spectral measurements into previously defined clusters through similarities observed in absorbance spectra. The classification system of the sampled tissue volume of the subjects is herein expanded to include classification of instrumentation variations.

Referring now to Figure 1, a general method of classifying spectral measurements into previously defined clusters is represented. Briefly the steps of the method are:

1. Measurement. (11)

2. Classification (12), in which the measured spectrum is assigned membership in any one of a number of predefined clusters 13.

3. Calibration (14), in which calibration models suited to each cluster are provided.

4. Identifying outliers (15)

MEASUREMENT

In general, instrumental variations may affect spectral response by producing either signal intensity changes, bandwidth changes, wavelength changes, or combinations thereof. These instrumental variations may include:

1. wavelength shifts;
2. nonlinear wavelength shifts;
3. wavelength expansions;
4. wavelength contractions;
5. nonlinear wavelength expansions;
6. source intensity drifts;
7. blackbody profile changes;
8. bandwidth changes;
9. resolution changes;
10. baseline deviations;
11. change with time;
12. temperature effects;
13. detector response;
14. differences in optical components (e.g. long-pass filters or fiber optics);
15. variation related to mounting of references;
16. differences in the optical interface to the sample (fiber spacing);
17. linearity;
18. detector cut-off;

and many others, which will be apparent to those skilled in the art. Spectra used for classification will typically be those of commonly known standards. Standards particularly useful for classifying shifts observed in the wavelength axis include polystyrene, rare earth oxides: holmium oxide, erbium oxide or dysprosium oxide, for example; or combinations thereof. Standards such as the diffuse reflectance standards supplied by Labsphere, Inc. (North Sutton NH) may be used to classify shifts in the intensity axis. Spectra of samples may also be used for classification. Additional intensity and wavelength standards will be known to those skilled in the art.

In the specific case of near-IR noninvasive glucose determination these standards cover the near-IR spectral region. Additional spectroscopy and chromatography-spectroscopy hyphenated techniques such as AA (atomic absorption spectroscopy) or GC-MS (gas chromatography mass spectroscopy) will each require their own standards which are known by those skilled in the art. Additionally, spectra of tissue phantoms collected on an instrument may also be used to characterize that instrument. Tissue phantoms are helpful in characterizing an instrument because spectra of these phantoms simulate noninvasive spectra of living tissue. There are tissue phantoms that simulate the absorption coefficients of various skin tissues; while others simulate the scattering coefficients of the body. Some common simulants include:

1. milk and milk products
2. milk products in combination with India ink, used to adjust the absorption coefficient;
3. emulsions of fatty substances in water maintained in solution with an emulsifier such as lecithin. One such commercial product is INTRALIPID, supplied by Kabivitrum AB (Stockholm, Sweden);
4. Intra-serum and Intra-gel. See K. Hazen, J. Welch, S. Malin, T. Ruchti, A Lorenz, T. Troy, S. Thenadil, T. Blank, *Intra-serum and intra-gel for modeling human skin tissue*, U.S. Patent Application Ser. No. 09/502,877 (February 10, 2000).

Other scattering and absorbing mediums are known to those skilled in the art. Concentrations of these tissue simulants may be adjusted to match the scattering and absorption coefficients of body tissues such as skin or internal organs.

It will also be apparent that a set of exemplar measurements is required for cluster definition and development of calibration models, in addition to the actual spectral measurements utilized for classification.

CLASSIFICATION

New spectral measurements are passed to a pattern classification system that classifies the measurements into previously defined clusters having a high degree of internal constistency through spectral features related to instrumental variation.

The classification itself includes the following steps:

1. Feature extraction; and

2. Classification of features according to a classification model.

FEATURE EXTRACTION

The process of feature extraction is more fully described in the parent application to the current application, U.S. Patent Application Ser. No. 09/359,191, *supra*. Feature extraction is any mathematical transformation that enhances a particular aspect or quality of the data that is useful for interpretation. Features may be of two categories:

1. Abstract, and
2. Simple.

Instruments may be classified by abstract features, meaning that they may be classified using computational methods. These methods may be supervised or unsupervised. Examples include plotting primary principal components versus one another and identifying clusters of results; discriminant analysis, such as measurement of the Mahalanobis distance, and k-means clustering. Additional methods will be readily apparent to those skilled in the art.

It is important to note that the clustering techniques listed above are not mutually exclusive. Clustering of raw spectra into one or more groups may be achieved through one or more of these methods and with combinations of these approaches.

Simple features are derived from an *a priori* understanding of the system, and can be related directly to an instrument parameter or component (or parameters or components). For example, the measured bandwith, noise characteristics, or linearity and detector cutoffs.

CLUSTER DEFINITION

As indicated above, clusters must have been previously defined, using a data set of exemplar spectral measurements. Cluster definition is the assignment of the measurements in the exploratory data set to clusters. After cluster definition, the measurements and class assignments are used to determine the mapping from the features to cluster assignment.

Cluster definition is performed using either supervised or unsupervised methods. In the supervised case, clusters may be defined using specific knowledge of the system. For example, source intensity and detector

temperatures have specific effects on the spectra, as demonstrated further below. The use of *a priori* information in this manner is the first step in supervised pattern recognition, which develops classification models when the class assignment is known.

Alternatively, clusters may be defined in an unsupervised manner using abstract features such as clustering within plots of principal component scores 'x' versus spectral loading 'y'. The result is that within a given cluster, all of the spectra have the same characteristics (interferences, instrument variations or sample issues). Clusters formed from features with physical meaning can be interpreted based on the known underlying phenomenon causing variation in the feature space.

CLASSIFICATION

Subsequent to class definition, a classifier is designed through supervised pattern recognition. A model is created, based on cluster definitions, that transforms a measured set of features to an estimated classification. The classification model is a method for determining a set of similarity measures with the predefined clusters. A decision rule assigns membership on the basis of a set of measures calculated by a decision engine.

CALIBRATION

Once the spectra have been classified into clusters, calibration models 14 are required for each cluster. Provision of the clusters is by one of two methods. In the preferred embodiment, a master calibration is developed for a first cluster, subsequently the master calibration is transferred to slave calibrations, one for each remaining cluster. An alternate embodiment, described further below, individually calculates calibrations for each cluster.

In the case of spectra collected utilizing a master and slave instrument, the term "calibration transfer," as commonly used in the art, may have different meanings. Calibration transfer may refer to transforming the slave spectral to look like the master spectra or *vice versa*. Additionally, spectra from both the master and the slave can be transferred to a common standard spectrum not present in the master or slave data sets. Furthermore, calibration transfer can refer to preprocessing steps, multivariate adjustments of the sample spectra, or adjustments to the predicted analyte concentrations based upon standards.

Referring now to Figure 2, the process of calibration transfer is shown. Calibrations 14 are required for each of *n* clusters. A master calibration 21 is

calculated for a first cluster; then, to provide calibrations for each of the remaining *n* clusters, the master calibration is transferred to slave calibrations. In this case, Calibration transfer refers to the process of transforming spectra on a slave instrument to match the characteristics of the master such that the master calibration can be applied to the slave spectra. For clusters, calibration transfer refers to the process of transforming spectra on a slave instrument to match the characteristics of the master such that a new calibration is generated that satisfies the specification of a cluster other than the one for which the master cluster was developed. The clusters may also be organized into groups of clusters, so that the master calibration is transferred to slave calibrations 23, which in turn are transferred to the slave calibration for the various clusters within each group.

Prior art methods of calibration transfer have been unsuccessful at modeling the complexity involved in providing calibration models for large numbers of instruments. Classifying the spectral measurements in to clusters having a high degree of internal consistency reduces the complexity of the problem to a level that makes it possible to apply calibration transfer to large numbers of instruments. Clustering the acquired spectra into a limited number of subgroups allows the calibration transfer issues to be broken up into subsets so that only a subset of the overall issues needs to be addressed between any two clusters. The result is that within a given cluster, all of the spectra have the same characteristics. A calibration is generated for a given cluster using spectra of samples containing the limited number of characteristics of that cluster. This eliminates having to deal with all of the variations possible in the raw measurements, significantly reducing the complexity that needs to be modeled by the calibration. Thus, the calibration utilized for a given cluster will have instrument variations included in the model that are similar to variations resulting from the analyzer employed. As each cluster has its own specification, the calibration transfer technique need only deal with the differences between those two clusters. For example, if the only difference between the clusters is a linear x-axis shift, then the calibration transfer technique need only deal with that parameter. This allows a more specific calibration transfer technique to be employed that is more robust, resulting in fewer factors, since fewer instrument variations need to be modeled. This will result in analysis of lower concentration analytes due to retention of signal to noise in the calibration transfer step.

Typically, calibration transfer techniques have to address instrument-to-instrument variations, such as wavelength shifts, nonlinear wavelength shifts, wavelength expansions (contraction), nonlinear wavelength expansions, source intensity

drifts, blackbody profile changes, bandwidth (resolution) changes, baseline deviations, change with time, temperature effects and others known to those skilled in the art.

However, in addition to instrumentation issues in calibration transfer, sample composition and sample handling considerations are very important in calibration transfer. Therefore, every sample analyzed for prediction of results should be an interpolation within the matrix space of the calibration data set; otherwise the calibration may not calculate an accurate prediction of analyte concentration. For example, if the calibration data set contains glucose samples ranging from 40 to 400 mg/dL, then predictions on samples with glucose outside of this range are suspect.

Thus, an important additional benefit of the invented classification algorithm is that outlier spectra may be identified, as shown in Figure 1. Each cluster has its own set of classification requirements. If a spectrum does not fall within the parameters of a given cluster, another cluster must be found which has parameters allowing analysis of that spectrum. If no cluster is found, then the sample is reported as an outlier rather than having a value reported for the analyte that may not be correct. In this manner, analysis on samples or instruments for which no satisfactory calibration model has been developed is prevented.

Spectra classified as outliers may assume two forms. In some instances, the outlier spectrum is close to a given cluster. In such a case, conventional calibration transfer techniques may be applied to this spectrum to transform the spectrum such that it falls within one of the clusters for which calibrations exist. In case of failure, the spectrum remains classified as an outlier, as will spectra that are characterized as gross outliers.

Classification and calibration transfer need not be limited to differences in instrumentation; it may also be applied to sample spectra for analyte prediction. For example, a cluster may be defined for healthy 18 to 30 year old men. A noninvasive glucose model may be built for this narrow demographic. A separate cluster may be for 18 to 30 year old non-pregnant women. Differences between the two clusters such as body fat and sampled volume may be addressed and a calibration transfer technique may be optimized for those differences based on relatively few spectra, thus allowing a glucose model initially generated for a narrow demographic to be gradually expanded.

The foregoing embodiment does not require calibrations to be built for every cluster, and thereby provides the important advantage of sparing the considerable time, money and effort required for building a separate calibration for every cluster. The actual calibration transfer technique employed may be selected from the many algorithms known to those skilled in the art.

An alternative embodiment classifies spectra into clusters, as in the preferred embodiment, and builds individual calibrations for each cluster. The requirement for calibration transfer is eliminated at the expense of the requirement for more spectra. However, an additional measure of reliability is gained by having clusters with well-defined and narrow parameters. As in the previous embodiment, fewer instrument variations need to be modeled so that early factors can focus on analyte information, resulting in a more robust model.

While the foregoing description of the invention has been directed to characterizing different instruments, the invented method also finds application in classifying spectra from a single instrument according to known variations in instrument components or known environmental variations. Two examples are described.

1. The cutoff of the 2.6 Ì m InGaAs detector employed in many near-IR analyzers blue shifts as the temperature of the detector decreases. Air spectra collected on a NICOLET 860, supplied by the Nicolet Instrument Corporation (Madison WI) with the InGaAs detector ranging from 0 to −20°C are presented in Figure 3A. The spectral cutoff region from 2400 to 2600 nm is expanded in Figure 3B. The cut-off decreases as the detector temperature decreases. Using reference or air spectra, a simple analysis, such as determining the wavelength at which 10% of the peak intensity is observed, allows the instrument to be classified as to whether the detector is properly cooled and to the extent that it was cooled.

2. As a tungsten halogen source increases in temperature, the blackbody emission increases in magnitude and additional light throughput is observed. For example, in diffuse reflectance spectra of a 5% Labsphere diffuse reflectance standard, the overall intensity is observed to increase as the source temperature increases, as shown in Figure 4. This effect can be used to classify the instrument in terms of the source temperature and total light throughput.

It is a relatively simple task to characterize the instrument in terms of additional components. Examples are slit width, affecting bandwidth or room temperature, affecting total light throughput. There are, however, a limited number of components to a spectrometer; it has been observed that instruments group into a relatively small number of clusters.

Once an instrument, or a given spectrum, is classified into a cluster, the appropriate calibration routine can be employed. A single analyzer may be loaded with multiple calibration routines, as the instrument may vary during its lifetime. In fact, environmental factors may cause the appropriate calibration to change with every spectrum collected.

This instrument classification approach was designed for use on noninvasive glucose analyzers where glucose is measured using diffuse reflectance spectra of skin on the human body. However, the same technology can be applied to any form of noninvasive analysis including but not limited to: noninvasive analysis of albumin, globulin, urea, creatinine, oxygen, hemoglobin A_1C and electrolytes such as Na^+, K^+, Cl^-. The technology may also be utilized in biomedical applications.

The classification approach disclosed here finds application in various fields of endeavor, the agricultural field, for example. Exemplary agricultural applications are: analysis of fat in milk, protein or moisture analysis in wheat, or analysis of sugars in fruits. The classification approach may also benefit pharmaceutical companies in the analysis of intact tablets or in the characterization of raw materials. Finally, petrochemical companies may utilize the method in classification of fuels, fuel by-products and in raw fuel materials. In general, this technique is useful wherever a large number of analyzers are utilized to quantify analytes in samples.

Although the invention is described herein with reference to the preferred embodiment, one skilled in the art will readily appreciate that other applications may be substituted for those set forth herein without departing from the spirit and scope of the present invention. Accordingly, the invention should only be limited by the claims included below.

CLAIMS

What is claimed is:

1. A method of characterizing spectrometer instruments according to instrumental variation, comprising the steps of:

providing standard spectral measurements from at least one spectrometer instrument; and

classifying said spectral measurements into predefined clusters on the basis of extracted spectral features; and

providing calibration models for each of said predefined clusters, wherein said calibration model compensates for said instrumental variation.

2. The method of Claim 1 wherein said instrumental variation comprises any of:

wavelength shifts;
nonlinear wavelength shifts;
wavelength expansions;
wavelength contractions;
nonlinear wavelength expansions;
source intensity drifts;
blackbody profile changes;
bandwidth changes;
resolution changes;
baseline deviations;
changes over time;
temperature effects;
detector response;
differences in optical components (e.g. long-pass filters or fiber optics);
variation related to mounting of references;
differences in the optical interface to the sample (fiber spacing);
linearity; and
detector cut-off.

3. The method of Claim 1, wherein said standard spectra are measured on a plurality of spectrometer instruments.

4. The method of Claim 1, wherein said standard spectral are measured on a single spectrometer instruments at successive time intervals.

5. The method of Claim 1, wherein said classifying step comprises the steps of:

extracting features; and

classifying said features according to a classification model and decision rule.

6. The method of Claim 5, wherein said feature extraction step comprises any mathematical transformation that enhances a particular aspect or quality of data that is useful for interpretation.

7. The method of Claim 3, wherein said classification model comprises means for determining a set of similarity measures with predefined classes.

8. The method of Claim 5, wherein said decision rule comprises means for assigning class membership on the basis of a set of measures calculated by a decision engine.

9. The method of Claim 4, wherein individual features are divided into two categories, said categories comprising:

abstract wherein said features are extracted using various computational methods ; and

simple features that are derived from an *a priori* understanding of a system, wherein said feature is directly related to an instrument parameter or component.

10. The method of Claim 7, wherein said abstract features are calculated using any of:

plotting primary principal components versus one another and identifying resulting clusters;

discrminant anlysis; and

k-means clustering.

11. The method of Claim 5, wherein said classification step further comprises the step of employing factor-based methods to build a model capable of representing variation in a measured spectrum related to variations in spectral response;

wherein projection of a measured absorbance spectrum onto said model constitutes a feature that represents spectral variation related to instrument variation.

12. The method of Claim 5, wherein said classifying step further comprises the steps of:

measuring the similarity of a feature to predefined clusters; and

assigning membership in a cluster.

13. The method of Claim 5, further comprising the step of:

assigning measurements in an exploratory data set to clusters.

14. The method of Claim 13, further comprising the step of:

using measurements and class assignments to determine a mapping from features to cluster assignments.

15. The method of Claim 13, further comprising the steps of:

defining clusters from said features in a supervised manner, wherein each set of features is divided into two or more regions, and wherein classes are defined by combinations of feature divisions;

designing a classifier subsequent to class definition through supervised pattern recognition by determining an optimal mapping or transformation from the feature space to a class estimate which minimizes the number of misclassifications; and

creating a model based on class definitions which transforms a measured set of features to an estimated classification.

16. The method of Claim 1, further comprising the step of providing calibration models for analysis of new sample measurements.

17. The method of Claim 16, wherein said calibration models model differences between said predefined clusters.

18. The method of Claim 16, wherein a master calibration model is developed for a first of said clusters from a set of exemplar spectra with reference values and pre-assigned classification definitions.

19. The method of Claim 18, further comprising the step of transferring said master calibration model to a plurality of slave calibration models, wherein a slave

BEST AVAILABLE COPY

calibration model is calculated for each remaining cluster, and wherein a transform modifies said master calibration model to a slave calibration model in accordance with principal features defining each of said classes.

20. The method of Claim 19, wherein said transferring step comprises the steps of:

transferring said master calibration model to a first slave calibration model;

transferring said first slave calibration model to a second slave calibration model;

and repeating said transfer from one slave calibration model to another slave calibration model, until a calibration has been provided for each of said predefined clusters;

wherein a transform modifies said transferred calibration models in accordance with principal features defining each of said clusters.

21. The method of Claim 18, further comprising the step of transferring said master calibration model to a plurality of slave calibration models, wherein a slave calibration model is calculated for each remaining cluster, and wherein a transform modifies said slave calibration model to said master calibration model in accordance with principal features defining each of said classes.

22. The method of Claim 21, wherein said transferring step comprises the steps of:

transferring said master calibration model to a first slave calibration model;

transferring said first slave calibration model to a second slave calibration model;

and repeating said transfer from one slave calibration model to another slave calibration model, until a calibration has been provided for each of said predefined clusters;

wherein a transform modifies said transferred calibration models in accordance with principal features defining each of said clusters.

23. The method of Claim 16, wherein a different calibration model is developed for each class, and wherein said calibration models are developed from a set of exemplar spectra, with reference values and pre-assigned cluster definitions.

24. The method of Claim 23, wherein a spectrum is assigned to one of many predefined clusters for which a calibration model has been developed.

25. The method of Claim 1, further comprising the steps of:

providing new spectral measurements;

comparing said new spectral measurements to each of said pre-defined clusters according to extracted spectral features;

reporting those measurements as outliers for which a matching cluster is not found.

26. A method of developing calibration models for spectral analysis comprising the steps of:

defining clusters from an exemplar data set of spectral measurements, wherein said clusters exhibit a high degree of internal similarity;

mapping said clusters to one another, wherein principal features distinguishing clusters from one another are determined;

calculating a calibration model for a first of said clusters, said calibration model comprising a master calibration;

transferring said master calibration to at least one slave calibration, wherein a slave calibration comprises a calibration derived by applying a transform to slave spectra such that the master calibration now models the difference between the master cluster and another cluster corresponding to said slave spectra.

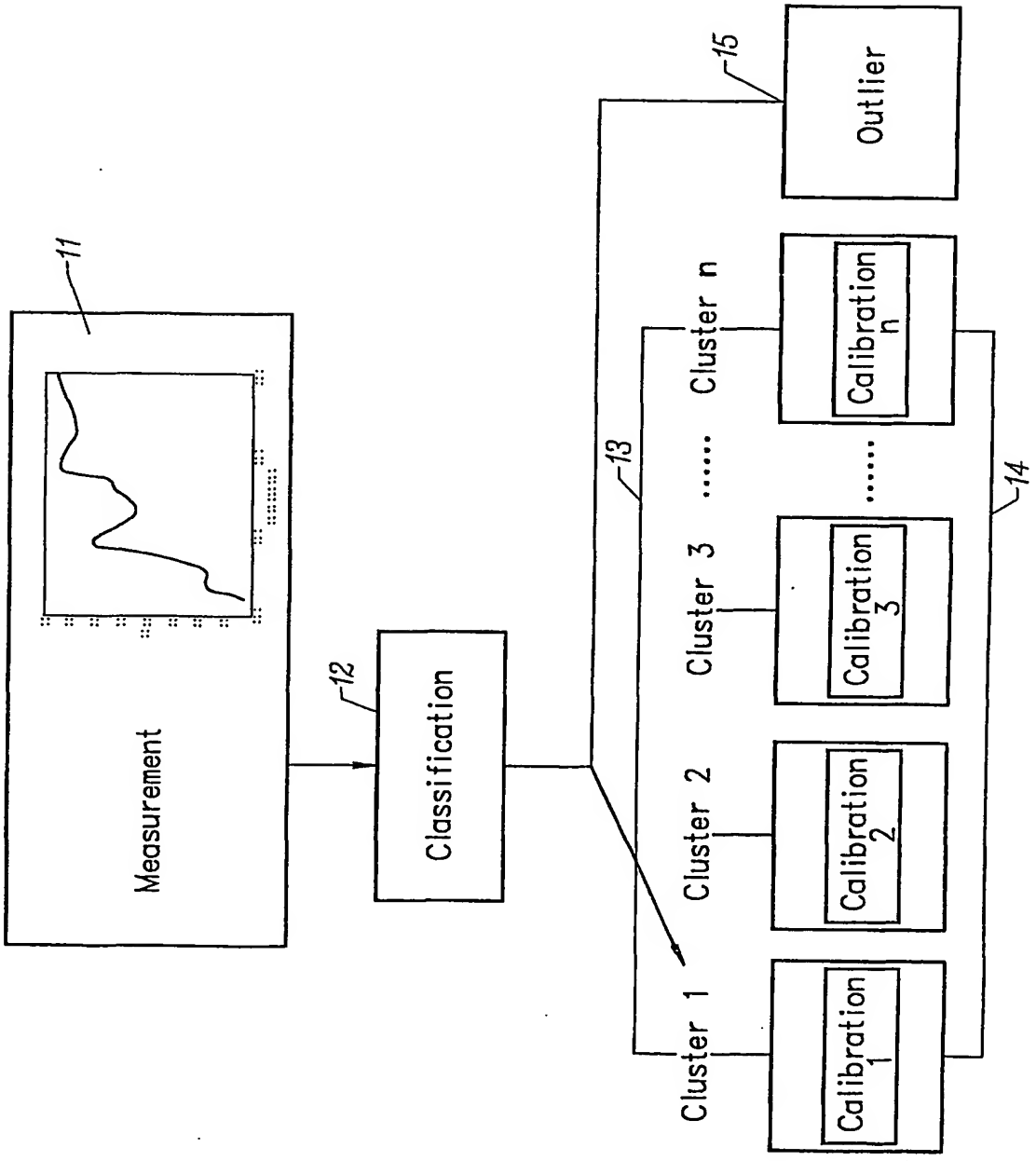

Measurement
11
Classification
12
13
14
15
Outlier
Cluster 1
Cluster 2
Cluster 3
Cluster n
Calibration 1
Calibration 2
Calibration 3
Calibration n

FIG. 1

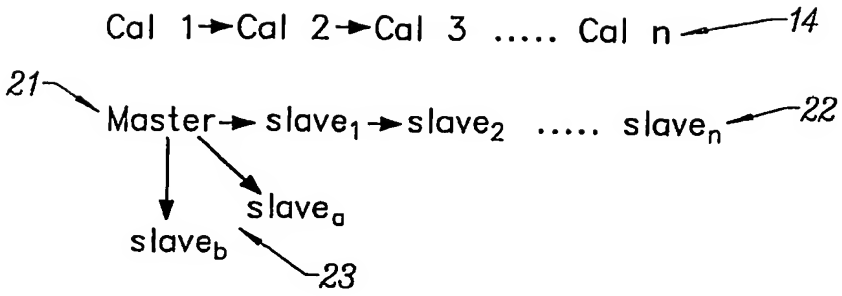
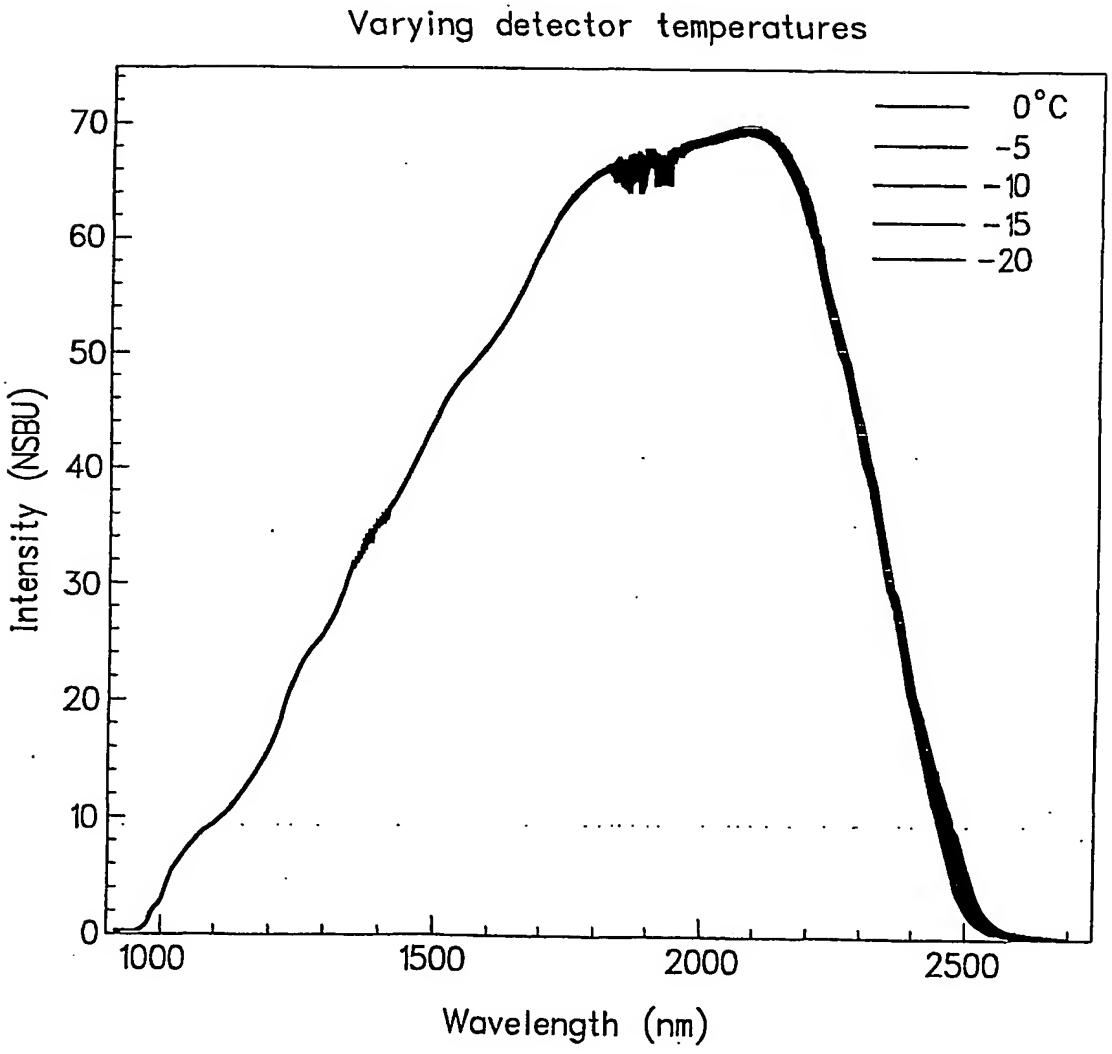


FIG. 2



BEST AVAILABLE COPY

FIG. 3A

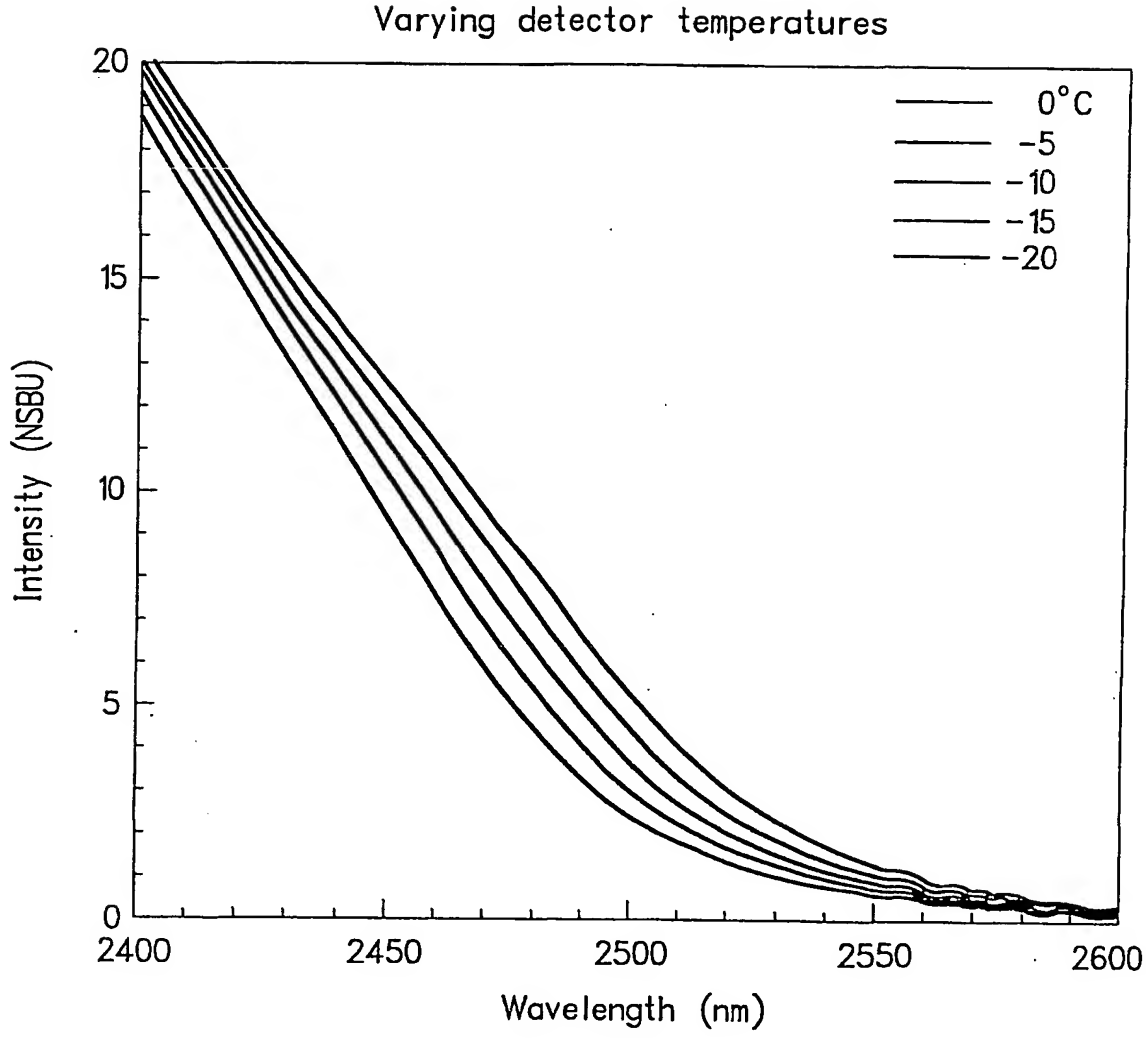

Varying detector temperatures
0°C
-5
-10
-15
-20
Intensity (NSBU)
20
15
10
5
0
2400
2450
2500
2550
2600
Wavelength (nm)

FIG. 3B

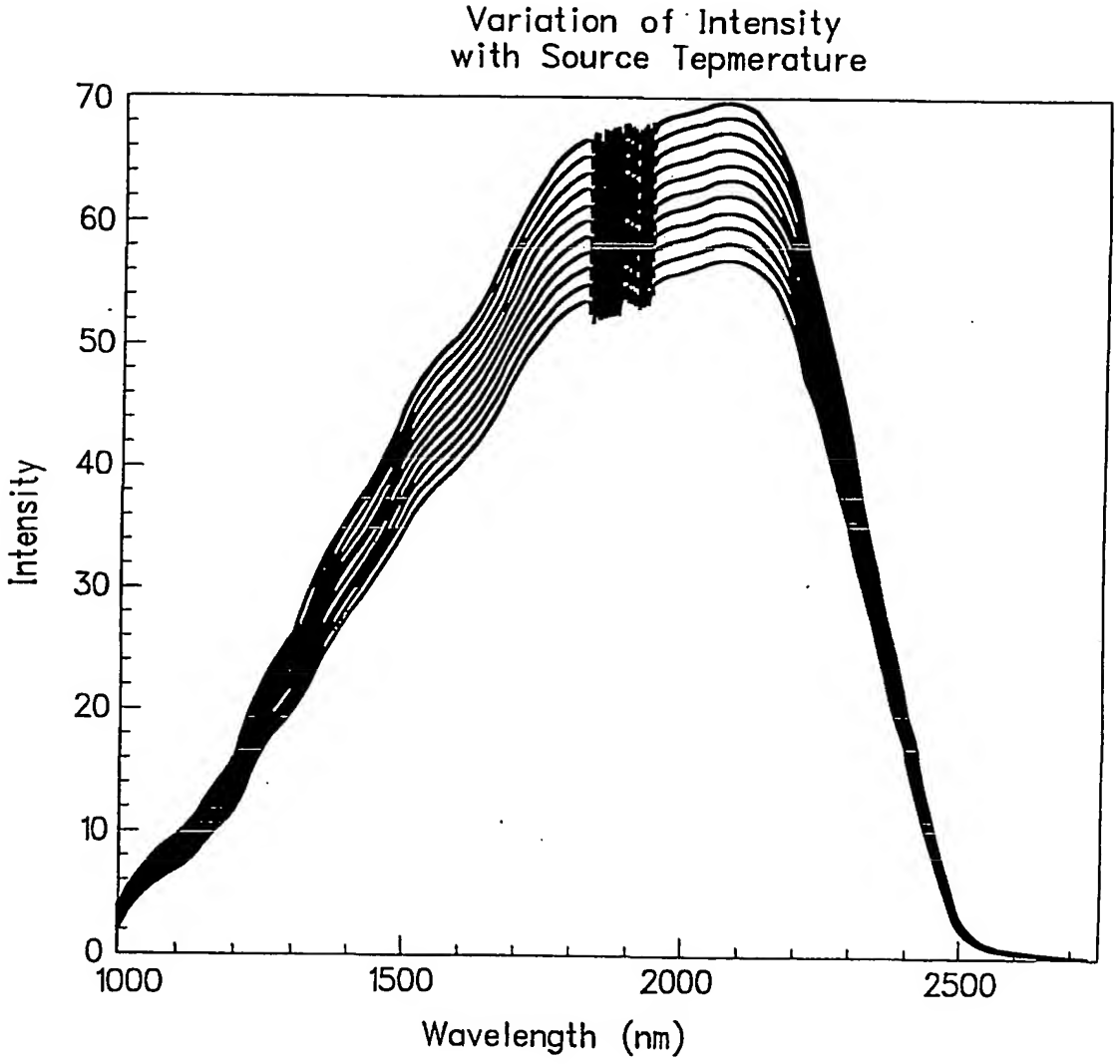


FIG. 4

BEST AVAILABLE COPY

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
28 March 2002 (28.03.2002)

PCT

(10) International Publication Number
WO 02/25233 A3

(51) International Patent Classification[7]: G01N 21/27

(21) International Application Number: PCT/US01/25706

(22) International Filing Date: 16 August 2001 (16.08.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
09/664,973 18 September 2000 (18.09.2000) US

(71) Applicant: INSTRUMENTATION METRICS, INC. [US/US]; 7470 West Chandler Blvd., Chandler, AZ 85226 (US).

(72) Inventors: HAZEN, Kevin, H.; 1534 W. Islandia Drive, Gilbert, AZ 85233 (US). BLANK, Thomas, B.; 2922 E. Tulsa Street, Chandler, AZ 85225 (US). MONFRE, Stephen, L.; 1289 E. Palo Blanco Way, Gilbert, AZ 85296 (US). RUCHTI, Timothy, L.; 1501 West Sea Haze Drive, Gilbert, AZ 85233 (US).

(74) Agents: GLENN, Michael et al.; Glenn Patent Group, Ste. L., 3475 Edison Way, Menlo Park, CA 94025 (US).

(81) Designated States *(national)*: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.

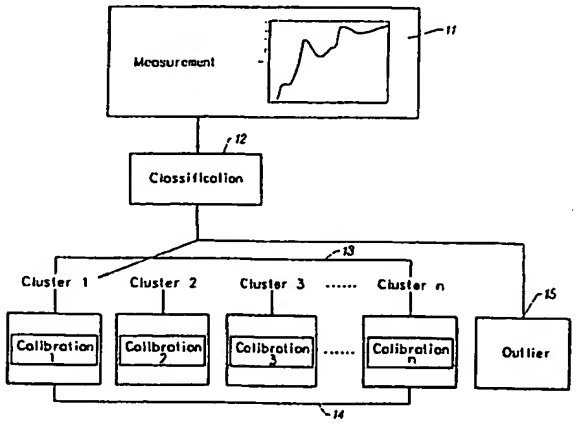
(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

(88) Date of publication of the international search report:
27 June 2002

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: METHOD OF CHARACTERIZING SPECTROMETERS AND PROVIDING CALIBRATION MODELS



(57) Abstract: Spectrometer instruments are characterized by classifying their spectra into previously defined clusters. The spectra are mapped to the clusters and a classification (12) is made based on similarity of extracted spectral features to one of the previously defined clusters. Calibration models for each cluster are provided to compensate for instrumental variation. Calibration models are provided either by transferring a master calibration (21) to slave calibrations (23) or by calculating a separate calibration for each cluster. A simplified method of calibration transfer maps clusters to each other, so that a calibration transferred between clusters models only the difference between the two clusters, substantially reducing the complexity of the model.

WO 02/25233 A3

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 01/25706

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 G01N21/27

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 G01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
EPO-Internal, WPI Data, PAJ, COMPENDEX, INSPEC, IBM-TDB

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 5 204 532 A (ROSENTHAL ROBERT D) 20 April 1993 (1993-04-20) column 3, line 61 -column 4, line 57	1,26
X	US 5 708 593 A (SABY CLAUDE-ALAIN ET AL) 13 January 1998 (1998-01-13) the whole document	1,26
X	US 5 710 713 A (WRIGHT LARRY G ET AL) 20 January 1998 (1998-01-20) column 5, line 28 - line 40	1,26
X	EP 0 768 522 A (ELF ANTAR FRANCE) 16 April 1997 (1997-04-16) the whole document	1,26

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
13 March 2002	04/04/2002

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Clevorn, J

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 01/25706

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	BOUVERESSE E, HARTMANN C, MASSART D L, LAST I R, PREBBLE K A: "Standardization of Near-Infrared Spectrometric Instruments" ANALYTICAL CHEMISTRY, vol. 68, no. 6, 15 March 1996 (1996-03-15), pages 982-990, XP002193004 cited in the application the whole document	1,26
A	WO 00 42907 A (INSTRUMENTATION METRICS INC) 27 July 2000 (2000-07-27) cited in the application the whole document	1,26

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/US 01/25706

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 5204532 A	20-04-1993	US 5028787 A	02-07-1991
		US 5086229 A	04-02-1992
		US 5077476 A	31-12-1991
		US 5068536 A	26-11-1991
		AU 2251292 A	12-01-1993
		CA 2111868 A1	23-12-1992
		EP 0590077 A1	06-04-1994
		JP 6508440 T	22-09-1994
		MX 9202953 A1	01-02-1993
		US 5576544 A	19-11-1996
		WO 9222804 A1	23-12-1992
		US 6066847 A	23-05-2000
		AT 163844 T	15-03-1998
		AU 5042190 A	13-08-1990
		CA 2045599 A1	20-07-1990
		DE 69032126 D1	16-04-1998
		DE 69032126 T2	15-10-1998
		EP 0456716 A1	21-11-1991
		ES 2114860 T3	16-06-1998
		JP 2191434 A	27-07-1990
		JP 5058735 B	27-08-1993
		KR 145695 B1	01-08-1998
		WO 9007905 A1	26-07-1990
		US 5218207 A	08-06-1993
		US 5365066 A	15-11-1994
		AT 145988 T	15-12-1996
		AU 8238791 A	23-01-1992
		CA 2086019 A1	28-12-1991
		DE 69123448 D1	16-01-1997
		DE 69123448 T2	22-05-1997
		EP 0536304 A1	14-04-1993
		IE 912231 A1	01-01-1992
		JP 5508336 T	25-11-1993
		NZ 238717 A	26-08-1994
		WO 9200513 A1	09-01-1992
		US 5362966 A	08-11-1994
		US 5436455 A	25-07-1995
		US 5438201 A	01-08-1995
		US 5574283 A	12-11-1996
		US 5237178 A	17-08-1993
		ZA 9104977 A	29-04-1992
US 5708593 A	13-01-1998	FR 2734360 A1	22-11-1996
		CA 2176915 A1	20-11-1996
		EP 0743512 A1	20-11-1996
		IL 118246 A	14-07-1999
		JP 8320239 A	03-12-1996
		NO 961956 A	20-11-1996
US 5710713 A	20-01-1998	US 5545895 A	13-08-1996
		AU 4928096 A	08-10-1996
		DE 19681255 T0	12-02-1998
		GB 2313915 A	10-12-1997
		WO 9629590 A1	26-09-1996
EP 0768522 A	16-04-1997	FR 2739928 A1	18-04-1997
		CA 2187943 A1	17-04-1997
		EP 0768522 A2	16-04-1997

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/US 01/25706

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
EP 0768522 A		IL	119427 A	31-01-2000
		JP	9178756 A	11-07-1997
		US	6128544 A	03-10-2000
WO 0042907 A	27-07-2000	US	6280381 B1	28-08-2001
		AU	2620900 A	07-08-2000
		BR	0008186 A	06-11-2001
		EP	1143850 A1	17-10-2001
		WO	0042907 A1	27-07-2000
		US	2001021803 A1	13-09-2001